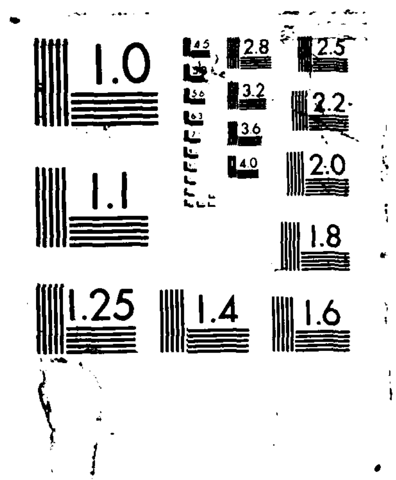
AD-A190 057 INFERENCE FOR A NONLINEAR SEMIMARTINGALE REGRESSION MODEL(U) FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS I W MCKEAGUE ET AL. NOV 87 FSU-TR-M-770 1/1

UNCLASSIFIED DAAL03-86-K-0094 F/G 12/3 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

ARO 23699.14-MA

DTIC FILE COPY

(2)

INFERENCE FOR A NONLINEAR SEMIMARTINGALE REGRESSION MODEL

AD-A190 857

by

Ian W. McKeague[1] and Klaus J. Utikal[2]

FSU Technical Report No. M-770
USARO Technical Report No. D-98
AFOSR Technical Report No. 87-213

November, 1987

Department of Statistics
The Florida State University
Tallahassee, FL 32306-3033

[1] Research supported by the Army Research Office under Grant DAAL03-86-K-0094.

[2] Research supported by the Air Force Office of Scientific Research under Contract F49620-85-C-0007

AMS 1980 subject classifications. 62M99, 62J02, 62G05.

Key words and phrases. Nelson-Aalen estimator, martingale central limit theorem, semimartingales, counting processes, diffusion processes, censored survival data.

88 1 27 064

INFERENCE FOR A NONLINEAR SEMIMARTINGALE REGRESSION MODEL

Ian W. McKeague and Klaus J. Utikal
Florida State University

Abstract

~~Consider~~ the semimartingale regression model

$$X(t) = X(0) + \int_0^t Y(s)\,\alpha(s, Z(s))\,ds + M(t),$$

where Y, Z are observable covariate processes, α is a (deterministic) function of both, time and the covariate process Z, and M is a square integrable martingale. Under the assumption that i.i.d. copies of X, Y, Z are observed continuously over a finite time interval, inference for the function $\alpha(t, z)$ is investigated. An estimator $\hat{A}$ for the time integrated $\alpha(t, z)$ and a kernel estimator of $\alpha(t, z)$ itself are introduced. For X a counting process, $\hat{A}$ reduces to the Nelson-Aalen estimator when Z is not present in the model. Various forms of consistency are proved, rates of convergence and asymptotic distributions of the estimators are derived. Asymptotic confidence bands for the time integrated $\alpha(t, z)$ and a Kolmogorov-Smirnov-type test of equality of α at different levels of the covariate are given.

1. Introduction.

A useful way of modelling the dependence of a counting process N on a covariate process Z was given by Aalen (1975). In his multiplicative intensity model N is supposed to have an intensity λ given by

$$\lambda(t) = \alpha(t)\, Z(t),$$

where α is an unknown, deterministic function of time (the hazard function). In the present paper we study inference for counting processes with intensities having general dependence on a covariate process Z, as in

$$\lambda(t) = Y(t)\, \alpha(t, Z(t)), \tag{1.1}$$

where α is an unknown, deterministic function of *both* time and the covariate process Z. The covariate process Y is taken to be an indicator process, assuming the value 1 when the counting process is under observation, zero otherwise.

An important example of our model arises in survival analysis. Suppose that the conditional hazard function $h(t\,|Z)$ for the survival time T of an individual given the covariate process Z has the form $h(t\,|Z) = \alpha(t, Z(t)\,)$. The observable portion of the individual's lifetime is given by $\tilde{T} = \min(T, C)$, where C is a censoring time. We observe $\tilde{T}$, $\delta = I(T \le C)$ and $Z(t)$ for $t \le \tilde{T}$. Let $N(t) = I(\tilde{T} \le t, \delta = 1)$, the counting process with a single jump at an uncensored survival time. If T and C are conditionally independent given Z then N has intensity (1.1), where $Y(t) = I(\tilde{T} \ge t)$ is the indicator that the individual is "at risk" at time t. We shall introduce an estimator $\hat{A}$ for the time-integrated conditional hazard function $A(\cdot, z) = \int_0^{\cdot} \alpha(s, z)\, ds$ which, in the special case of a time-independent covariate Z, coincides with an estimator proposed in unpublished work of Beran (1981). Dabrowska (1987) recently obtained a weak convergence result for Beran's estimator by proving a "conditional" analogue of Theorem 4 of Breslow and Crowley (1974). We obtain asymptotic results for our estimator by using a martingale approach (in particular, Rebolledo's martingale central limit theorem) which enables us to give quite simple proofs and to avoid the restrictive assumption of time-independent covariate Z.

For another example of our model, consider a pure jump process describing the motion of a particle on a finite state space $\{1, 2, ..., m\}$. Let the intensity $\alpha_{ij}(t, s)$ of transition from state i to state j depend on the (calendar) time t and on the time s spent in state i since the last jump. Let $Y_i(t)$ be the indicator that the particle is in state i at time t. Then the counting process $N_{ij}(t)$ which registers the number of transitions from state i to state j up to time t has intensity

$$\lambda(t) = Y_i(t-)\, \alpha_{ij}(t, L(t-)), \tag{1.2}$$

where $L(t)$ is the length of time which at time t has elapsed since the last jump. In the terminology of Markov renewal processes (Pyke,1961), $L(t)$ is the backward recurrence time. In the case that

α_{ij} only depends on calendar time t, inference for α_{ij} has been studied by Aalen (1975, 1978). In the case that α_{ij} only depends on the backward recurrence time $L(t)$, N_{ij} is a Markov renewal process for which inference has been studied by Gill (1980).

The results of the paper will be developed for the nonlinear semimartingale regression model

$$X(t) = X(0) + \int_0^t Y(t)\, \alpha(s, Z(s))ds + M(t), \tag{1.3}$$

where M is a square integrable martingale and Y, Z, α are as before. This includes diffusion processes as well as the counting processes mentioned above. An estimator $\hat{A}(t,z)$ of the time-integrated conditional "hazard" function

$$A(t,z) = \int_0^t \alpha(s,z)\, ds$$

will be introduced. For counting processes our estimator $\hat{A}$ coincides with the Nelson-Aalen estimator if the covariate process Z is constant. A kernel estimator $\hat{\alpha}$ of α will be obtained from $\hat{A}$, as was done by Ramlau-Hansen (1983) for the Nelson-Aalen estimator.

The estimators $\hat{\alpha}$ and $\hat{A}$ are defined in Section 2. Consistency and asymptotic distribution results for $\hat{\alpha}$ and $\hat{A}$ are given in Section 3. In Section 4 we derive confidence bands for $A(\cdot, z)$ at any fixed level z of the covariate and introduce a Kolmogorov-Smirnov type statistic for testing the hypothesis that $A(\cdot, z_1)$ and $A(\cdot, z_2)$ coincide (equivalently $\alpha(\cdot, z_1)$ and $\alpha(\cdot, z_2)$ coincide) at different levels z_1, z_2. Technical lemmas used in the proofs of the main results are given in Section 5.

2. The Estimators.

$(\Omega, \mathcal{F}, P)$ will denote a complete probability space and $(\mathcal{F}_t, t \in [0,1])$ a nondecreasing right-continuous family of sub-σ-fields of $\mathcal{F}$ such that $\mathcal{F}_0$ contains all P-null sets in $\mathcal{F}$. All processes are indexed by $t \in [0,1]$. The process $M = (M(t), \mathcal{F}_t)$ is assumed to be a square integrable martingale with mean zero and paths which are right-continuous on $[0,1)$ with left limits on $(0,1]$. Suppose that the quadratic characteristic $\langle M \rangle$ of M has the form

$$\langle M \rangle(t) = \int_0^t \gamma(s, Z(s), Y(s))\, ds, \tag{2.1}$$

where γ is a bounded, measurable function. The covariate processes Y and Z are assumed to be predictable and Y is an indicator process. For simplicity, Z is supposed to be scalar valued. We assume that the processes X, Y, Z and M are related by the equation (1.3) which can be written in differential form

$$dX(t) = Y(t)\, \alpha(t, Z(t))\, dt + dM(t), \tag{2.2}$$

where α is a bounded Borel function. In the counting process case

$$\gamma(t, Z(t), Y(t)) = Y(t)\,\alpha(t, Z(t)).$$

In the diffusion process case (without censoring) we have $Y(t) \equiv 1,\ Z(t) = X(t)$,

$$M(t) = \int_0^t \sigma(s, X(s))\,dW(s),$$

$$\gamma(t, Z(t), Y(t)) = \sigma^2(t, X(t)),$$

where W is a Wiener process, $\sigma^2(t,z)$ is the infinitesimal variance of the diffusion and $\alpha(t,z)$ is the drift function.

In order to define the estimators $\hat{\alpha}$ and $\hat{A}$ we need the following notation. For $z \in \mathbf{R}$, I_z denotes an interval of length w_n containing z, where $w_n \to 0$ as $n \to \infty$. Let $(X_i, Y_i, Z_i, M_i), i = 1, \ldots, n$ denote n independent copies of the generic processes X, Y, Z, M which satisfy model (2.1), (2.2). Assume that X_i and Y_i are observable continuously over the time interval [0,1] and $Z_i(t)$ is observable at least when $Y_i(t) \neq 0$. Define

$$X^{(n)}(t,z) = \sum_{i=1}^{n} \int_0^t I\{Z_i(s) \in I_z\} Y_i(s)\,dX_i(s), \tag{2.3}$$

$$Y^{(n)}(t,z) = \sum_{i=1}^{n} I\{Z_i(t) \in I_z\}\,Y_i(t). \tag{2.4}$$

As an estimator of A we propose

$$\hat{A}(t,z) = \int_0^t \frac{1}{Y^{(n)}(s,z)}\,X^{(n)}(ds,z),$$

where $1/0 \equiv 0$. Also, for $t \in (0,1)$ set

$$\hat{\alpha}(t,z) = \frac{1}{b_n} \int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,\hat{A}(ds,z),$$

where K is a bounded, nonnegative kernel function with compact support, integral 1 and $b_n > 0$ is a bandwidth parameter, $b_n \to 0$.

We note that for the above estimation of A at a fixed z the processes X, Z only need to be observed at times when Z belongs to the neighborhood I_z of z. We can show that $\hat{A}$ and $\hat{\alpha}$ yield asymptotically well behaved estimators of A and α when we shrink I_z (i.e. let $w_n \to 0$) and let $b_n \to 0$ at appropriate rates as the sample size increases. If K is left continuous and of bounded variation, then by integration by parts (see Dellacherie and Meyer (1982), Chapter VIII, (19.4)) for n sufficiently large

$$\frac{1}{b_n} \int_O^1 K\Big(\frac{t-s}{b_n}\Big)\,\hat{A}(ds,z) = \frac{1}{b_n} \int_0^1 \hat{A}(s-,z)\,dK\Big(\frac{t-s}{b_n}\Big)$$

almost surely for all $t \in [t_1, t_2]$, where $0 < t_1 < t_2 < 1$. Thus, for n sufficiently large, we can choose a Lebesgue-Stieltjes version of the process $(\hat{\alpha}(t,z),\, t \in [t_1, t_2])$:

$$\hat{\alpha}(t,z) = \frac{1}{b_n} \int_0^1 \hat{A}(s-, z)\, dK\Big(\frac{t-s}{b_n}\Big). \tag{2.5}$$

This version of $\hat{\alpha}$ is used in Theorem 2(c).

3. Main Results.

We shall consider estimation of $A(t,z)$, $\alpha(t,z)$ over $0 \le t \le 1$, $0 \le z \le 1$. Let C be a set in $\mathbf{R}$ containing $\cup_{z \in [0,1]} I_z^{(n)}$ for some $n \ge 1$. The following assumptions are supposed to hold for all (t,z) belonging to $[0,1] \times C$.

(A1) For each t, the random vector $(Z(t), Y(t))$ is absolutely continuous with respect to the product of the Lebesgue and counting measure. Denote the corresponding density by $f_{Z(t)Y(t)}(z,y)$. Also, suppose that for fixed z, y this density is integrable in t.

(A2) $f_{Z(t)Y(t)}(z,1)$ is bounded away from zero.

(A3) $f_{Z(t)Y(t)}(z,1)$ is continuous as a function of t and z for each fixed y.

(B1) α, γ are continuous functions of t and z for each fixed y.

(B2) α is Lipschitz, i.e. there exists a constant K such that

$$|\alpha(t_1, z_1) - \alpha(t_2, z_2)| \le K \|(t_1 - t_2, z_1 - z_2)\|$$

for all t_1, t_2, z_1, z_2, where $\|\cdot\|$ denotes the Euclidian norm on $\mathbf{R}^2$.

THEOREM 1. (a) Suppose that A1-A3, B1 hold and $nw_n \to \infty$ as $n \to \infty$. Then

$$\sup_z E \sup_t |\hat{A}(t,z) - A(t,z)|^2 \to 0 \tag{3.1}$$

as $n \to \infty$.

(b) Suppose, in addition, that B2 holds and $nw_n^3 \to 0$ as $n \to \infty$. Then

$$nw_n E\, |\hat{A}(t,z) - A(t,z)|^2 \to \int_0^t h(s,z)\, ds \tag{3.2}$$

uniformly over $(t,z) \in [0,1]^2$ as $n \to \infty$ and

$$\limsup_{n \to \infty} nw_n E \sup_t |\hat{A}(t,z) - A(t,z)|^2 \le 4 \int_0^1 h(s,z)\, ds \tag{3.3}$$

uniformly over $z \in [0,1]$ as $n \to \infty$, where

$$h(s,z) = \frac{\gamma(s,z,1)}{f_{Z(s)Y(s)}(z,1)}. \tag{3.4}$$

Proof. Define

$$M^{(n)}(t,z) = \sum_{i=1}^{n} \int_0^t I\{Z_i(s) \in I_z\}\, Y_i(s)\, dM_i(s)$$

$$\alpha^{(n)}(t,z) = \sum_{i=1}^{n} I\{Z_i(t) \in I_z\}\, Y_i(t)\, \alpha(t, Z_i(t))$$

$$\gamma^{(n)}(t,z) = \sum_{i=1}^{n} I\{Z_i(t) \in I_z\}\, Y_i(t)\, \gamma(t, Z_i(t), Y_i(t)).$$

It follows from (2.1)-(2.4) that

$$X^{(n)}(dt,z) = \alpha^{(n)}(t,z)\, dt + M^{(n)}(dt,z)$$

$$d\langle M^{(n)}(\cdot\,,z)\rangle(t) = \gamma^{(n)}(t,z)\, dt.$$

The Doob-Meyer decomposition of $\hat{A}$ is

$$\hat{A}(t,z) = \int_0^t \frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)}\, ds + \int_0^t \frac{1}{Y^{(n)}(s,z)}\, M^{(n)}(ds,z). \tag{3.5}$$

Therefore

$$|\hat{A}(t,z) - A(t,z)|^2 = I_1(t) + I_2(t) + I_3(t),$$

where

$$I_1(t) = \Big(\int_0^t \frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)}\, ds - \int_0^t \alpha(s,z)\, ds\Big)^2$$

$$I_2(t) = \Big(\int_0^t \frac{1}{Y^{(n)}(s,z)}\, M^{(n)}(ds,z)\Big)^2$$

$$I_3(t) = 2I(t_1)I(t_2).$$

Now

$$E\sup_t I_1(t) \le \int_0^1 E\left|\frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)} - \alpha(s,z)\right|^2 ds$$

and by Lemmas 4 and 5

$$E\left[\frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)} - \alpha(s,z)\right]^2 = o(1)$$

uniformly in s, z as $n \to \infty$ if α is continuous. Also

$$nw_n E\left[\frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)} - \alpha(s,z)\right]^2 = o(1) \tag{3.6}$$

uniformly in s, z as $n \to \infty$ if α is Lipschitz and $n\,w_n^3 \to 0$. Next, by Doob's inequality

$$nw_n E \sup_t I_2(t) \le 4nw_n E\left[\int_0^1 \frac{1}{Y^{(n)}(s,z)}\, M^{(n)}(ds,z)\right]^2 = 4nw_n \int_0^1 E\,\frac{\gamma^{(n)}(s,z)}{\left(Y^{(n)}(s,z)\right)^2}\,ds,$$

where the r.h.s. of the last equality tends to $4\int_0^1 h(s,z)ds$ uniformly in z as $n \to \infty$ by Lemma 6. This proves (3.1),(3.3). To show (3.2) we observe that

$$nw_n E\, I_2(t) = nw_n E\left[\int_0^t \frac{1}{Y^{(n)}(s,z)}\, M^{(n)}(ds,z)\right]^2 = nw_n \int_0^1 E\,\frac{\gamma^{(n)}(s,z)}{\left(Y^{(n)}(s,z)\right)^2}\,ds.$$

This completes the proof.

The next result can be viewed as the analogue of Theorem 1 for the estimator $\hat{\alpha}$.

THEOREM 2. (a) Suppose that A1-A3, B1 hold and $b_n \sim w_n$, $nw_n^2 \to \infty$ as $n \to \infty$. Then

$$E\left[\hat{\alpha}(t,z) - \alpha(t,z)\right]^2 \to 0$$

for every $t \in (0,1)$ uniformly in z as $n \to \infty$ and

$$\int_0^1 E\left[\hat{\alpha}(t,z) - \alpha(t,z)\right]^2 dt \to 0$$

uniformly in z as $n \to \infty$.

(b) Suppose, in addition, that B2 is satisfied and $nw_n^4 \to 0$ as $n \to \infty$. Then

$$nw_n^2 E\left[\hat{\alpha}(t,z) - \alpha(t,z)\right]^2 \to \kappa h(s,z)$$

for every $t \in (0,1)$ uniformly in z as $n \to \infty$ and

$$nw_n^2 \int_0^1 E\left[\hat{\alpha}(t,z) - \alpha(t,z)\right]^2 dt \to \kappa \int_0^1 h(t,z)dt$$

uniformly in z as $n \to \infty$, where $\kappa = \int_{-\infty}^{\infty} K^2(u)\,du$.

(c) Suppose A1-A3, B1, B2 hold, K is left continuous, of bounded variation and $nw_n \to \infty$, $nw_n^3 \to 0$, $nw_n b_n^2 \to \infty$. Let $0 < t_1 < t_2 < 1$. Then, for the version of $\hat{\alpha}$ given by (2.5),

$$\sup_z E\Big(\sup_{t\in[t_1,t_2]} |\hat{\alpha}(t,z) - \alpha(t,z)|\Big) = O\Big(\frac{1}{\sqrt{nw_n}\,b_n}\Big) + O(b_n).$$

Proof. From the decomposition (3.5) it follows that

$$[\hat{\alpha}(t,z) - \alpha(t,z)]^2 = [I_1(t)]^2 + [I_2(t)]^2 + 2[I_1(t)I_2(t)],$$

where

$$I_1(t) = \frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\frac{1}{Y^{(n)}(s,z)}\,M^{(n)}(ds,z)$$

$$I_2(t) = \frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)}\,ds - \alpha(t,z).$$

Now

$$nw_n^2 E\,[I_1(t)]^2 = nw_n^2\,E\int_0^1 \Big(\frac{1}{b_n}K\Big(\frac{t-s}{b_n}\Big)\Big)^2\,\frac{\gamma^{(n)}(s,z)}{(Y^{(n)}(s,z))^2}\,ds.$$

Under the hypothesis of part (a) it follows from Lemma 1 and Lemma 3 that $nw_n^2 E\,[I_1(t)]^2$ is uniformly bounded, hence $E\,[I_1(t)]^2 \to 0$ and $\int_0^1 E\,[I_1(t)]^2\,dt \to 0$ as $n \to \infty$. Under the hypothesis of part (b) it follows from Lemma 8 that $nw_n^2 E\,[I_1(t)]^2 \to \kappa\,h(t,z)$ for all $t \in (0,1)$ uniformly in z as $n \to \infty$. The bounded convergence theorem implies that $nw_n^2 \int_0^1 E\,[I_1(t)]^2\,dt \to \kappa \int_0^1 h(t,z)\,dt$ uniformly in z as $n \to \infty$. By the Cauchy-Schwarz inequality, $E\,|I_1(t)I_2(t)| \le \{E\,[I_1(t)]^2\,E\,[I_2(t)]^2\,\}^{1/2}$, the proof of parts (a), (b) will be complete if we we show that

$$E\,[I_2(t)]^2 = \begin{cases} o(1) & \text{under the conditions of part (a) for all } t \in (0,1) \text{ uniformly in } z \\ O(w_n^2) & \text{under the conditions of part (b) for all } t \in (0,1) \text{ uniformly in } z \end{cases}$$

and similarly for $\int_0^1 E\,[I_2(t)]^2\,dt$. This is done in Lemma 7. To prove (c) we observe that $\hat{\alpha}(t,z) - \alpha(t,z) = I_3(t,z) + I_4(t,z)$, where

$$I_3(t,z) = \frac{1}{b_n}\int_0^1 (\hat{A}(s-,z) - A(s-,z))\,dK\Big(\frac{t-s}{b_n}\Big)$$

$$I_4(t,z) = \frac{1}{b_n}\int_0^1 A(s-,z)\,dK\Big(\frac{t-s}{b_n}\Big) - \alpha(t,z).$$

Now

$$|I_3(t,z)| \le (2/b_n)\,V(K)\sup_{0\le s\le 1}|\hat{A}(s-,z) - A(s-,z)|,$$

where $V(K)$ is the total variation of K. Thus, by Theorem 1(b)

$$\sup_z E\sup_{0\le t\le 1}|I_3(t,z)| = O\Big(\frac{1}{\sqrt{nw_n}\,b_n}\Big).$$

By integration by parts, for n sufficiently large

$$I_4(t,z) = \frac{1}{b_n}\int_0^1 K\left(\frac{t-s}{b_n}\right) A(ds,z) - \alpha(t,z) = \frac{1}{b_n}\int_0^1 K\left(\frac{t-s}{b_n}\right)\alpha(s,z)\,ds - \alpha(t,z).$$

Therefore, by B2, $|I_4(t,z)| = O(b_n)$ uniformly in $(t,z) \in [t_1,t_2] \times [0,1]$. This completes the proof of the theorem.

REMARK. The conditions on w_n, b_n of part (c) of Theorem 2 are satisfied for the sequences $b_n = n^{-\beta}$ and $w_n = n^{-\delta}$, where $1/3 < \delta < 1$ and $0 < \beta < (1-\delta)/2$.

Asymptotic distribution results for $\hat{A}$ will be established under additional assumptions on X that will make functional central limit theorems for martingales easily applicable to the martingale part of (3.5). In the sequel $D[0,1]$ denotes the space of real valued functions on $[0,1]$ which are right continuous on [0,1) and whose left limits exist on $(0,1]$, equipped with the Skorohod topology. Also, $D[0,1]^l$ denotes the product space of ι copies of $D[0,1]$. For an account of weak convergence in $D[0,1]$ we refer to Billingsley (1968).

THEOREM 3. Suppose that X has continuous sample paths or is a counting process, A1-A3, B1-B2 hold and $nw_n \to \infty$, $nw_n^3 \to 0$ as $n \to \infty$. Then for $z_1, ..., z_l \in [0,1]$, all distinct, the process

$$(\sqrt{nw_n}(\hat{A}(t,z_r) - A(t,z_r)), t \in [0,1])_{r=1}^{l}$$

converges weakly in $D[0,1]^l$ as $n \to \infty$ to the Gaussian process

$$(U(t,z_r), t \in [0,1])_{r=1}^{l}$$

with zero mean and covariance function

$$\mathrm{Cov}(U(t_1,z_{r_1}), U(t_2,z_{r_2})) = \delta_{r_1,r_2}\int_0^{t_1 \wedge t_2} h(s,z_{r_1})\,ds$$

(where δ_{r_1,r_2} denotes the Kronecker symbol).

Proof. By decomposition (3.5)

$$\sqrt{nw_n}(\hat{A}(t,z_r) - A(t,z_r)) = \sqrt{nw_n}(I_3(t) + I_4(t)),$$

where

$$I_3(t) = \int_0^t \frac{\alpha^{(n)}(s,z_r)}{Y^{(n)}(s,z_r)}\,ds - \int_0^t \alpha(s,z_r)\,ds,$$

$$I_4(t) = \int_0^t \frac{1}{Y^{(n)}(s,z_r)}\,M^{(n)}(ds,z_r).$$

Observe that

$$nw_n E \sup_t (I_3(t))^2 \leq nw_n \int_0^1 E\left[\frac{\alpha^{(n)}(s,z_r)}{Y^{(n)}(s,z_r)} - \alpha(s,z_r)\right]^2 ds \to 0$$

as $n \to \infty$ by (3.6). Therefore it will be sufficient to show that

$$(\sqrt{nw_n} I_4(t))_{r=1}^l \to (U(t,z_r))_{r=1}^l \tag{3.7}$$

weakly in $D[0,1]^l$ as $n \to \infty$. If X has continuous sample paths, so does the square integrable martingale $I_4(t)$ and by Liptser and Shiryayev (1980) it will be sufficient to show that

$$nw_n \langle I_4 \rangle(t) \xrightarrow{P} \int_0^t h(s,z)\, ds \tag{3.8}$$

for all t, z_r as $n \to \infty$. Since

$$\left\langle \int_0^\cdot \frac{1}{Y^{(n)}(s,z_r)} M^{(n)}(ds,z_r) \right\rangle(t) = \int_0^t \frac{\gamma^{(n)}(s,z_r)}{(Y^{(n)}(s,z_r))^2}\, ds,$$

(3.8) follows from Lemma 6. If X is a counting process, by Rebolledo (1978) we will have to verify, in addition, the Lindeberg condition

$$nw_n \int_0^1 \frac{\gamma^{(n)}(s,z_r)}{(Y^{(n)}(s,z_r))^2} I\{\sqrt{nw_n}\frac{1}{Y^{(n)}(s,z_r)} > \epsilon\}\, ds \xrightarrow{P} 0,$$

for all $\epsilon > 0$. But by application of the Cauchy-Schwarz inequality

$$E\, nw_n \frac{\gamma^{(n)}(s,z_r)}{(Y^{(n)}(s,z_r))^2} I\{\sqrt{nw_n}\frac{1}{Y^{(n)}(s,z_r)} > \epsilon\}$$

$$\leq \left\{E\left[nw_n \frac{1}{Y^{(n)}(s,z_r)}\right]^4\right\}^{1/2} \left\{E\left[\frac{\gamma^{(n)}(s,z_r)}{nw_n}\right]^4 P[\sqrt{nw_n}\frac{1}{Y^{(n)}(s,z_r)} > \epsilon]\right\}^{1/4}.$$

By Corollary 1 and Lemma 3 the r.h.s. above tends to zero for all z_r if for all s, z_r

$$P[\sqrt{nw_n}\frac{1}{Y^{(n)}(s,z_r)} > \epsilon] \to 0.$$

This follows from Lemma 3 and the Markov inequality. So far we have only shown weak convergence in $D[0,1]$ for each z_r. Note that $(I_4(t))_{r=1}^l$ is a vector of square integrable martingales that are orthogonal whenever $I_{z_{r_1}} \cap I_{z_{r_2}} = \emptyset$, which is true for sufficiently large n. Therefore (3.7) follows from the previous arguments by application of the Cramér-Wold device.

The next theorem gives the asymptotic finite dimensional distributions of the estimator $\hat{\alpha}$.

THEOREM 4. Suppose that X has continuous sample paths or is a counting process, A1-A3, B1, B2 hold and $b_n \sim w_n$, $nw_n^2 \to \infty$, $nw_n^4 \to 0$ as $n \to \infty$. Then for all $z_1, \ldots, z_l \in [0,1], t_1, \ldots, t_k \in (0,1)$, all distinct,

$$(\sqrt{n}w_n(\hat{\alpha}(t_j,z_r)-\alpha(t_j,z_r)))_{j=1,r=1}^{k,l}$$

converges in distribution to the Gaussian random array $(V(t_j,z_r))_{j=1,r=1}^{k,l}$ with mean zero and covariance

$$\mathrm{Cov}(V(t_{j_1},z_{r_1}),V(t_{j_2},z_{r_2}))=\delta_{j_1,j_2}\delta_{r_1,r_2}\,\kappa h(t_{j_1},z_{r_1}).$$

Proof. By decomposition (3.5)

$$\sqrt{n}w_n(\hat{\alpha}(t_j,z_r)-\alpha(t_j,z_r))=\sqrt{n}w_n(I_5+I_6),$$

where

$$I_5=\frac{1}{b_n}\int_0^1 K\Big(\frac{t_j-s}{b_n}\Big)\frac{\alpha^{(n)}(s,z_r)}{Y^{(n)}(s,z_r)}\,ds-\alpha(t_j,z_r),$$

$$I_6=\frac{1}{b_n}\int_0^1 K\Big(\frac{t_j-s}{b_n}\Big)\frac{1}{Y^{(n)}s,z_r)}\,M^{(n)}(ds,z_r).$$

It follows from Lemma 7 that $\sqrt{n}w_nI_5\xrightarrow{P}0$ as $n\to\infty$. Therefore it will be sufficient to show that

$$(\sqrt{n}w_nI_6)_{j=1,r=1}^{k\,l}\xrightarrow{\mathcal{D}}(V(t_j,z_r))_{j=1,r=1}^{k,l}.$$

Now $I_6=I_6(1)$, where

$$I_6(\tau)=\int_0^\tau\frac{1}{b_n}K\Big(\frac{t_j-s}{b_n}\Big)\frac{1}{Y^{(n)}(s,z_r)}\,M^{(n)}(ds,z_r),$$

and $(I_6(\tau))_{j=1,r=1}^{k,l}$ is an array of $k\,l$ square integrable martingales that are orthogonal for n sufficiently large. If X has continuous sample paths it follows from Remark 2 in Liptser and Shiryayev (1980) and the Cramér-Wold device that we only need to show $nw_n{}^2\langle I_6\rangle(1)\xrightarrow{P}\kappa h(t_j,z_r)$ as $n\to\infty$ for all t_j, z_r. This is done in Lemma 8. If X is a counting process we have to verify, in addition, that

$$nw_n^2\int_0^1\Big(\frac{1}{b_n}K\Big(\frac{t-s}{b_n}\Big)\Big)^2\frac{\gamma^{(n)}(s,z)}{(Y^{(n)}(s,z))^2}\,I\Big\{\sqrt{n}w_n\,\frac{1}{b_n}K\Big(\frac{t-s}{b_n}\Big)\frac{1}{Y^{(n)}(s,z)}>\epsilon\Big\}\,ds\xrightarrow{P}0$$

as $n\to\infty$ for all $\epsilon>0$. As in the proof of Theorem 3 this follows from

$$P[\sqrt{n}w_n\,\frac{1}{b_n}K\Big(\frac{t-s}{b_n}\Big)\frac{1}{Y^{(n)}(s,z)}>\epsilon]\le P[\sqrt{n}\,\frac{1}{Y^{(n)}(s,z)}>\tilde{\epsilon}]$$

(for some $\tilde{\epsilon}>0$, since K is bounded)

$$=P[nw_n\,\frac{1}{Y^{(n)}(s,z)}>\tilde{\epsilon}\sqrt{n}w_n]\to0$$

(by Lemma 3 and the Markov inequality).

This completes the proof of the theorem.

4. Confidence Bands and Hypothesis Tests.

In order to use the previous theorems for inference, an estimate of $h(t,z)$ (as defined in (3.4)) and of $\int_0^t h(s,z)\,ds = H(t,z)$ is needed. The following theorem provides consistent estimators for both of these quantities in the case that X is a counting process.

THEOREM 5. Suppose that A1-A3, B1, B2 hold, X is a counting process. Define

$$\hat{H}(t,z) = nw_n \int_0^t \frac{1}{\left(Y^{(n)}(s,z)\right)^2} X^{(n)}(ds,z),$$

for $t \in [0,1]$,

$$\hat{h}(t,z) = \frac{1}{b_n} \int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,\hat{H}(ds,z),$$

for $t \in (0,1)$. If $b_n \sim w_n$ $nw_n^2 \to \infty$ as $n \to \infty$ then

$$E\,|\,\hat{h}(t,z) - h(t,z)| \to 0$$

for all $t \in (0,1)$ uniformly in $z \in [0,1]$ as $n \to \infty$. If $nw_n \to \infty$ then

$$E \sup_t |\hat{H}(t,z) - H(t,z)| \to 0$$

uniformly in $z \in [0,1]$ as $n \to \infty$.

Proof. By decomposition (3.5)

$$|\hat{h}(t,z) - h(t,z)| \le I_1(t) + I_2(t),$$

where

$$I_1(t) = \Big|\,\frac{nw_n}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\frac{\alpha^{(n)}(s,z)}{(Y^{(n)}(s,z))^2}\,ds - h(t,z)\,\Big|,$$

$$I_2(t) = \Big|\,\frac{nw_n}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\frac{1}{(Y^{(n)}(s,z))^2}\,M^{(n)}(ds,z)\,\Big|.$$

By Lemma 8 (with K^2 replaced by κK) and $b_n \sim w_n$ we have

$$\sup_z E\, I_1(t) \to 0.$$

Next

$$E\,[I_2(t)]^2 \le \frac{1}{{b_n}^2}\int_0^1 K^2\Big(\frac{t-s}{b_n}\Big)\, n^2{w_n}^2\, E\,\frac{\gamma^{(n)}(s,z)}{\big(Y^{(n)}(s,z)\big)^4}\,ds \to 0$$

for all $t \in (0,1)$ uniformly in z as $n \to \infty$ by Lemmas 1, 3. Similarly

$$|\hat{H}(t,z) - H(t,z)| = \Big|\, nw_n \int_0^t \frac{1}{\big(Y^{(n)}(s,z)\big)^2}\, X^{(n)}(ds,z) - \int_0^t h(s,z)\,ds \,\Big| \le I_3(t) + I_4(t),$$

where

$$I_3(t) = \Big|\, nw_n \int_0^t \frac{\alpha^{(n)}(s,z)}{\big(Y^{(n)}(s,z)\big)^2}\,ds - \int_0^t h(s,z)\,ds\,\Big|,$$

$$I_4(t) = \Big|\, nw_n \int_0^t \frac{1}{\big(Y^{(n)}(s,z)\big)^2}\, M^{(n)}(ds,z)\,\Big|.$$

But

$$\sup_z E \sup_t I_3(t) \le \int_0^1 \sup_{z,s} E\,\Big|\, nw_n \frac{\alpha^{(n)}(s,z)}{\big(Y^{(n)}(s,z)\big)^2} - h(s,z)\,\Big|\, ds \to 0$$

by Lemma 6 (with $\gamma^{(n)}$ replaced by $\alpha^{(n)}$) and

$$\sup_z E \sup_t\, [I_4(t)]^2 \le \int_0^1 \sup_{z,s} E\,\Big[(nw_n)^2 \frac{\gamma^{(n)}(s,z)}{\big(Y^{(n)}(s,z)\big)^4}\Big]\, ds \to 0$$

as $n \to \infty$ by Doob's inequality and Lemmas 1, 3.

In the diffusion process case, in which $\sigma^2(t,z)$ is assumed to be known, the following theorem provides consistent estimators for $h(t,z)$ and $H(t,z)$.

THEOREM 6. Suppose A1-A3, B1, B2 hold, X is a diffusion process. Define

$$\hat{H}(t,z) = nw_n \int_0^t \frac{\sigma^2(s,z)}{Y^{(n)}(s,z)}\,ds$$

for $t \in [0,1]$,

$$\hat{h}(t,z) = \frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,\hat{H}(ds\;\;z)$$

for $t \in (0,1)$. If $b_n \sim w_n$, $nw_n^2 \to \infty$ then

$$E\,|\hat{h}(t,z) - h(t,z)| \to 0$$

for all $t \in (0,1)$ uniformly in $z \in [0,1]$ as $n \to \infty$. If $nw_n \to \infty$ then

$$E \sup_t |\hat{H}(t,z) - H(t,z)| \to 0$$

uniformly in $z \in [0,1]$ as $n \to \infty$.

Proof. It follows from $b_n \sim w_n$ and Lemma 8 (where we replace $\gamma^{(n)}(s,z)$ by $\sigma^2(s,z)\,Y^{(n)}(s,z)$ and K^2 by κK) that

$$E\,|\hat{h}(t,z) - h(t,z)| = E\left|\frac{nw_n}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,\frac{\sigma^2(s,z)}{Y^{(n)}(s,z)}\,ds - h(t,z)\right| \to 0$$

for all $t \in (0,1)$ uniformly in z. Next

$$E \sup_t |\hat{H}(t,z) - H(t,z)| \le \sup_{s,z} E\left|nw_n\frac{\sigma^2(s,z)}{Y^{(n)}(s,z)} - h(s,z)\right|,$$

which tends to zero by Lemma 6 with $\gamma^{(n)}(s,z)$ replaced by $\sigma^2(s,z)\,Y^{(n)}(s,z)$. This completes the proof of the theorem.

Confidence bands for $A(\cdot,z)$.

Under the conditions of Theorem 3

$$\sqrt{nw_n}\,\frac{\sqrt{H(1,z)}}{H(1,z)+H(t,z)}\Big(\hat{A}(t,z) - A(t,z)\Big) \to W^0\Big(\frac{H(t,z)}{H(1,z)+H(t,z)}\Big)$$

weakly in $D[0,1]$ as $n \to \infty$, where W^0 is the Brownian bridge process. By application of Theorem 5 in the counting process case or Theorem 6 in the diffusion process case we obtain the following asymptotic $100(1-\alpha)\%$ confidence band for $A(\cdot,z)$:

$$\hat{A}(t,z) \pm c_\alpha\sqrt{\frac{\hat{H}(1,z)}{nw_n}}\Big(1 + \frac{\hat{H}(t,z)}{\hat{H}(1,z)}\Big), \quad t \in [0,1],$$

where

$$P[\sup_{t\in[0,1/2]} |W^0(t)| > c_\alpha] = \alpha.$$

A table for the distribution of $\sup_{t\in[0,1/2]} |W^0(t)|$ can be found in Hall and Wellner (1980).

Testing equality of A at two different levels of the covariate.

We now introduce a test statistic for testing the null hypothesis $H_0 : A(t,z_1) = A(t,z_2)$ for all $t \in [0,1]$, where z_1, z_2 are two prechosen values of z. Define

$$A_{1\,2}(t) = A(t,z_1) - A(t,z_2)$$

$$\hat{A}_{1\,2}(t) = \hat{A}(t, z_1) - \hat{A}(t, z_2)$$

$$H_{1\,2}(t) = H(t, z_1) - H(t, z_2)$$

$$\hat{H}_{1\,2}(t) = \hat{H}(t, z_1) - \hat{H}(t, z_2).$$

Then under the conditions of Theorem 3 we have that

$$\sqrt{nw_n}\frac{\sqrt{H_{1\,2}(1)}}{H_{1\,2}(1) + H_{1\,2}(t)}(\hat{A}_{1\,2}(t) - A_{1\,2}(t)) \to W^0\Big(\frac{H_{1\,2}(t)}{H_{1\,2}(1) + H_{1\,2}(t)}\Big)$$

weakly in $D[0,1]$ as $n \to \infty$. Set

$$\hat{\xi} = \sqrt{nw_n \hat{H}_{1\,2}(1)} \sup_{t\in[0,1]}\left|\frac{\hat{A}_{1\,2}(t) - A_{1\,2}(t)}{\hat{H}_{1\,2}(1) + \hat{H}_{1\,2}(t)}\right|.$$

Then in the counting process and diffusion process cases considered above $\hat{\xi}\xrightarrow{\mathcal{D}}\xi$ as $n \to \infty$, where $\xi \stackrel{\mathcal{D}}{=} \sup_{t\in[0,1/2]} |W^0(t)|$. Therefore an asymptotic test of size α can be carried out by rejecting H_0 if and only if $\hat{\xi} > c_\alpha$, where $P(\xi > c_\alpha) = \alpha$. Finally we mention that Theorem 4 can be used to construct asymptotic χ^2-tests as in Rao (1973) for testing equality of α at any finite number of values of t and z.

5. Technical Lemmas.

LEMMA 1. Suppose that A1, A3, B1 hold and $nw_n \to \infty$ as $n \to \infty$. Then

$$E\left[\frac{1}{nw_n}\gamma^{(n)}(s,z)\right]^k = (g(s,z))^k + o(1) \tag{5.1}$$

for all nonnegative integers k, uniformly in s, z as $n \to \infty$, where

$$g(s,z) = f_{Z(s)\,Y(s)}(z,1)\,\gamma(s,z,1).$$

Proof. By the multinomial theorem

$$E\,[\gamma^{(n)}(s,z)]^k = \sum_{j_1+\cdots+j_n=k} \frac{k!}{j_1!\cdots j_n!}\prod_{i=1}^{n} E\,[I\{Z_i(s) \in I_z\}\,Y_i(s)\,\gamma(s, Z_i(s), Y_i(s))]^{j_i}.$$

Since $|Z_i(s) - z| < w_n$ implies $|\gamma(s, Z_i(s), 1) - \gamma(s, z, 1)| < \epsilon_n$ uniformly in s, z for some $\epsilon_n \to 0$ as $n \to \infty$, we have for $j_i \neq 0$

$$E\,[I\{Z_i(s) \in I_z\}\,Y_i(s)\,\gamma(s, Z_i(s), Y_i(s))]^{j_i} = E\,[I\{Z_i(s) \in I_z\}\,Y_i(s)\,(\gamma(s,z,1) + O(\epsilon_n)]^{j_i}$$

$$= (((\gamma(s,z,1))^{j_i} + O(\epsilon_n))\,E\,[I\{Z_i(s) \in I_z\}\,Y_i(s)] = I_1.$$

Also, by uniform continuity of f,

$$|E\left[I\{Z_i(s)\in I_z\}Y_i(s)\right]-w_nf_{Z(s)Y(s)}(z,1)|\le\int_{I_z}|f_{Z(s)Y(s)}(u,1)-f_{Z(s)Y(s)}(z,1)|\,du=w_n\,o(1)$$

uniformly in s and z. Therefore

$$I_1=(\gamma(s,z,1))^{j_i}w_nf_{Z(s)Y(s)}(z,1)+w_no(1) \tag{5.2}$$

uniformly in s, z and

$$E\left[\frac{\gamma^{(n)}(s,z)}{nw_n}\right]^k=\left(\frac{1}{nw_n}\right)^k\left\{k!\binom{n}{k}\left\{w_n(f_{Z(s)Y(s)}(z,1)\,\gamma(s,z,1)+o(1))\right\}^k\right.$$

$$+\sum_{l=1}^{k-1}k!\binom{n}{l}\binom{k-1}{l-1}O(w_n^l)$$

$$=(g(s,z))^k+o(1).$$

COROLLARY 1. Suppose that A1, A3 hold and $nw_n\to\infty$ as $n\to\infty$. Then

$$E\left[\frac{\gamma^{(n)}(s,z)}{nw_n}-g(s,z)\right]^k\to 0 \tag{5.3}$$

as $n\to\infty$ for all $s,z,k\ge 1$,

$$\mathrm{Var}\left[\frac{\gamma^{(n)}(s,z)}{nw_n}\right]\to 0 \tag{5.4}$$

uniformly in s,z as $n\to\infty$ and

$$\mathrm{Var}\left[\frac{\gamma^{(n)}(s,z)}{nw_n}\right]^2\to 0$$

uniformly in s,z as $n\to\infty$.

LEMMA 2. Suppose $X\sim$ binomial (n,p), $0<p\le 1$. Let

$$X^*=\begin{cases}1/X, & \text{if } X>0;\\ 0, & \text{if } X=0.\end{cases}$$

Then for each integer $k\ge 1$

$$E\left[X^*\right]^k\le\left(\frac{k+1}{np}\right)^k.$$

Proof.

$$E\,[X^*]^k = \sum_{i=1}^{n} \frac{1}{i^k}\,\frac{n!}{i!\,(n-i)!}\,p^i\,q^{n-i} = \sum_{i=1}^{n} \frac{1}{i^k}\,\frac{(i+1)\cdots(i+k)}{(i+k)!}\,\frac{n!}{(n-i)!}\,p^i\,q^{n-i}$$

$$\le \sum_{i=1}^{n} (k+1)^k \frac{n!}{(i+k)!\,(n-i)!}\,p^i\,q^{n-i} = \frac{(k+1)^k\,n!}{p^k\,(n+k)!} \sum_{i=1}^{n} \frac{(n+k)!}{(i+k)!\,(n-i)!} p^{i+k} q^{n-i}$$

$$\le \left(\frac{k+1}{n\,p}\right)^k.$$

LEMMA 3. Suppose A1, A2 hold and $nw_n \to \infty$ as $n \to \infty$. Then

$$E\left[\frac{nw_n}{Y^{(n)}(s,z)}\right]^k = O(1)$$

uniformly in s, z as $n \to \infty$ for every integer $k \ge 1$, where $1/0 \equiv 0$.

Proof. Set $m = \inf_{s,z} f_{Z(s)Y(s)}(z,1)$. Then $Y^{(n)}(s,z)$ has a binomial distribution with parameters $(n, p^{(n)}(s,z))$, where $p^{(n)}(s,z) \ge m w_n$, so the previous lemma applies.Therefore

$$E\left[\frac{nw_n}{Y^{(n)}(s,z)}\right]^k \le \left(\frac{(k+1)\,nw_n}{n\,mw_n}\right)^k.$$

In the following lemma we will use the notation $J^{(n)}(s,z) = I\{Y^{(n)}(s,z) \ne 0\}$.

LEMMA 4. Suppose that A1, A2 hold and $nw_n \to \infty$ as $n \to \infty$. Then

$$E\,|1 - J^{(n)}(s,z)|^k \le \exp\{-nw_n \inf_{s,z} f_{Z(s)Y(s)}(z,1)\}$$

for each integer $k \ge 1$.

Proof. Set $m = \inf_{s,z} f_{Z(s)Y(s)}(z,1)$. Then $m > 0$ and

$$E\,|1 - J^{(n)}(s,z)|^k = P\,(J^{(n)}(s,z) = 0) = (1 - P\,[Z(s) \in I_z,\, Y(s) = 1])^n$$

$$\le (1 - mw_n)^n \le \exp\{-m\,nw_n\}.$$

LEMMA 5.

$$\left|\frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)} - J^{(n)}(s,z)\,\alpha(s,z)\right| = \begin{cases} o(1) & \text{uniformly in } s,z \text{ if } \alpha \text{ is continuous} \\ O(w_n) & \text{uniformly in } s,z \text{ if } \alpha \text{ is Lipschitz.} \end{cases}$$

Proof. By definition of $\alpha^{(n)}(s,z)$

$$|\alpha^{(n)}(s,z) - \alpha(s,z)\,Y^{(n)}(s,z)| \le \epsilon^{(n)}\,Y^{(n)}(s,z),$$

where $\epsilon^{(n)} = o(1)\,(= O(w_n))$ uniformly is s,z if α is continuous (if α is Lipschitz).

LEMMA 6. Suppose that A1 - A3, B1 hold and $nw_n \to \infty$ as $n \to \infty$. Then

$$\sup_{s,z} E\left|nw_n \frac{\gamma^{(n)}(s,z)}{\left(Y^{(n)}(s,z)\right)^2} - h(s,z)\right| \to 0,$$

where

$$h(s,z) = \frac{\gamma(s,z,1)}{f_{Z(s)Y(s)}(z,1)}.$$

Proof.

$$E\left|nw_n \frac{\gamma^{(n)}(s,z)}{\left(Y^{(n)}(s,z)\right)^2} - h(s,z)\right| \le I_1 + I_2,$$

where

$$I_1 = E\left|nw_n \frac{\gamma^{(n)}(s,z)}{\left(Y^{(n)}(s,z)\right)^2} - J^{(n)}(s,z)\,h(s,z)\right|$$

$$I_2 = h(s,z)\,E\,|1 - J^{(n)}(s,z)|.$$

By Lemma 4, $I_2 \to 0$ uniformly in s,z as $n \to \infty$. Now by application of the Cauchy-Schwarz inequality

$$I_1 \le \left\{E\left[\frac{nw_n}{Y^{(n)}(s,z)}\right]^4\right\}^{1/2} \{I_3\}^{1/2},$$

where

$$I_3 = E\left[\frac{\gamma^{(n)}(s,z)}{nw_n} - \left(\frac{Y^{(n)}(s,z)}{nw_n}\right)^2 h(s,z)\right]^2.$$

Also $\frac{1}{2} I_3 \le I_4 + I_5$, where

$$I_4 = E\left[\frac{\gamma^{(n)}(s,z)}{nw_n} - g(s,z)\right]^2$$

$$I_5 = C^2 E\left[\left(\frac{Y^{(n)}(s,z)}{nw_n}\right)^2 - \left(f_{Z(s)Y(s)}(z,1)\right)^2\right]^2,$$

where $C = \sup_{s,z} h(s,z)$. But

$$I_4 = \mathrm{Var}\Big[\frac{\gamma^{(n)}(s,z)}{nw_n}\Big] + o(1)$$

uniformly in s, z as $n \to \infty$ by Lemma 1 and (5.4) and

$$I_5 = \mathrm{Var}\Big[\frac{Y^{(n)}(s,z)}{nw_n}\Big]^2 + o(1)$$

uniformly in s, z as $n \to \infty$ by Lemma 1 with $\gamma(s,z,y) \equiv 1$. Therefore $I_4 \to 0$ uniformly is s, z as $n \to \infty$ by Corollary 1, and $I_5 \to 0$ uniformly in s, z as $n \to \infty$ by Corollary 1 with $\gamma(s,z,y) \equiv 1$.

LEMMA 7. Suppose that A1, A2 hold and for some $\theta > 0$, $nw_n^{1+\theta} \to \infty$ as $n \to \infty$. Set

$$I(t) = E\left[\frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)}\,ds - \alpha(t,z)\right]^2.$$

Then

$$I(t) = \begin{cases} o(1) & \text{for all } t \in (0,1) \text{ uniformly in } z \text{ if } \alpha \text{ is continuous,} \\ O(w_n^2) & \text{for all } t \in (0,1) \text{ uniformly is } z \text{ if } \alpha \text{ is Lipschitz,} \end{cases}$$

and

$$\int_0^1 I(t)\,dt = \begin{cases} o(1) & \text{uniformly in } z \text{ if } \alpha \text{ is continuous,} \\ O(w_n^2) & \text{uniformly is } z \text{ if } \alpha \text{ is Lipschitz.} \end{cases}$$

Proof. $I(t) \le 3(I_1(t) + I_2(t) + I_3(t))$, where

$$I_1(t) = E\left[\frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\left|\frac{\alpha^{(n)}(s,z)}{Y^{(n)}(s,z)} - J^{(n)}(s,z)\,\alpha(s,z)\right| ds\right]^2$$

$$I_2(t) = E\left[\frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,(1 - J^{(n)}(s,z))\,\alpha(s,z)\,ds\right]^2$$

$$I_3(t) = \left(\frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,\alpha(s,z)\,ds - \alpha(t,z)\right)^2.$$

By Lemma 5,

$$I_1(t) = \begin{cases} o(1) & \text{uniformly in } t, z \text{ if } \alpha \text{ is continuous,} \\ O(w_n^2) & \text{uniformly in } t, z \text{ if } \alpha \text{ is Lipschitz,} \end{cases}$$

which implies

$$\int_0^1 I_1(t)\,dt = \begin{cases} o(1) & \text{uniformly in } z \text{ if } \alpha \text{ is continuous,} \\ O(w_n^2) & \text{uniformly in } z \text{ if } \alpha \text{ is Lipschitz.} \end{cases}$$

Next

$$I_2(t) \le \frac{1}{b_n}\int_0^1 K^2\Big(\frac{t-s}{b_n}\Big)\,ds\,\Big(\sup_{s,z}\alpha(s,z)\Big)^2 \sup_{s,z}\frac{1}{b_n}E\Big(1-J^{(n)}(s,z)\Big)^2.$$

But $\sup_{s,z} E\left(1-J^{(n)}(s,z)\right)^2 = O(1/(nw_n))^k$ for all integers $k \ge 1$ by Lemma 4. This and $b_n \sim w_n$ imply $w_n{}^{-2} I_2(t) = O(1/(nw_n^{1+3/k}))^k$ for $3/k < \theta$ uniformly in t, z as $n \to \infty$. Thus $I_2(t) = O(w_n^2)$ and $\int_0^1 I_2(t)\,dt = O(w_n^2)$. Finally

$$I_3(t) = \begin{cases} o(1) & \text{for all } t \in (0,1) \text{ uniformly in } z \text{ if } \alpha \text{ is coninuous,} \\ O(w_n^2) & \text{for all } t \in (0,1) \text{ uniformly in } z \text{ if } \alpha \text{ is Lipschitz.} \end{cases}$$

Uniform boundedness of $I_3(t)$ and the dominated convergence theorem imply

$$\int_0^1 I_3(t)\,dt = \begin{cases} o(1) & \text{uniformly in } z \text{ if } \alpha \text{ is continuous,} \\ O(w_n^2) & \text{uniformly in } z \text{ if } \alpha \text{ is Lipschitz.} \end{cases}$$

This proves the lemma.

LEMMA 8. Suppose that A1-A3, B1 hold and $nw_n^2 \to \infty$ as $n \to \infty$. Then for each $\tau \in [0,1]$

$$\sup_z E\left| nw_n^2 \int_0^\tau \Big(\frac{1}{b_n}K\Big(\frac{t-s}{b_n}\Big)\Big)^2 \frac{\gamma^{(n)}(s,z)}{\left(Y^{(n)}(s,z)\right)^2}\,ds - \kappa(\tau,t)\,h(t,z)\right| \to 0$$

as $n \to \infty$ for all $t \in (0,1)$, where

$$\kappa(\tau,t) = \begin{cases} 0, & \text{if } \tau < t; \\ \int_{-\infty}^0 K^2(u)\,du, & \text{if } \tau = t; \\ \kappa, & \text{if } \tau > t. \end{cases}$$

Proof. For $\tau < t$ the theorem is obvious. Suppose $\tau \ge t$. Then

$$E\left| nw_n^2 \int_0^\tau \Big(\frac{1}{b_n}K\Big(\frac{t-s}{b_n}\Big)\Big)^2 \frac{\gamma^{(n)}(s,z)}{(Y^{(n)}(s,z))^2}\,ds - \kappa(\tau,t)\,h(t,z)\right| \le I_1(t) + I_2(t),$$

where

$$I_1(t) = E\,\frac{w_n}{b_n^2}\int_0^\tau K^2\Big(\frac{t-s}{b_n}\Big)\left| nw_n \frac{\gamma^{(n)}(s,z)}{\left(Y^{(n)}(s,z)\right)^2} - h(s,z)\right| ds$$

$$I_2(t) = \left|\frac{w_n}{b_n^2}\int_0^\tau K^2\Big(\frac{t-s}{b_n}\Big)\,h(s,z)\,ds - \kappa(\tau,t)\,h(t,s)\right|.$$

Note that $w_n/b_n \to 1$. Therefore $I_1(t) \to 0$ for all $t \in (0,1)$ uniformly in z as $n \to \infty$ by Lemma 6 and $I_2(t) \to 0$ for all $t \in (0,1)$ uniformly in z as $n \to \infty$ by continuity of h.

References

Aalen, O.O. (1975). Statistical inference for a family of counting processes. Ph.D. dissertation, Univ. of California, Berkeley.

Aalen, O.O. (1978). Nonparametric inference for a family of counting processes. Ann. Statist. 6, 701-726.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Tech. Rep., University of California, Berkeley.

Billingsley, P. (1986). Convergence of Probability Measures. Wiley, New York.

Breslow, N. and Crowley, J.J. (1974). A large sample study of the life table and product limit estimates under random censorship. Ann. Statist. 2, 437-443.

Dabrowska, D.M. (1987) Nonparametric regression with censored survival time data. To appear in Scand. J. Statist.

Dellacherie, C. and Meyer, P.-A. (1982). Probabilities and Potential B. North Holland, Amsterdam.

Gill, R.D. (1980). Nonparametric estimation based on censored observations for a Markov renewal process. Z. Wahrscheinlichkeitsthe. verw. Geb. 53, 97-116.

Hall, W.J. and Wellner, J.A. (1980). Confidence bands for a survival curve with censored data. Biometrika 67, 133-143.

Liptser, R.S. and Shiryayev, A.N. (1980). A functional central limit theorem for semimartingales. Theor. Prob. Appl. 25, 667-688.

Pyke, R. (1961). Markov renewal processes: definitions and preliminary properties. Ann. Math. Statist. 32, 1231-1242.

Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. Ann. Statist. 11, 453-466.

Rao, C.R. (1973). Linear Statistical Inference and Its Applications. Wiley, New York.

Rebolledo, R. (1978). Sur les applications de la théorie des martingales à l'étude statistique d'une famille de processus ponctuels. Lect. Notes Math. 636, 27-70.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARO 23699.14-MA	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) Inference for a Nonlinear Semimartingale Regression Model		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER FSU Techinical Report M-770
7. AUTHOR(s) Ian W. McKeague and Klaus J. Utikal		8. CONTRACT OR GRANT NUMBER(s) DAAL03-86-K-0094
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University Department of Statistics Tallahassee, FL 32306-3033		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE November, 1987
		13. NUMBER OF PAGES 23
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

N/A

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nelson-Aalen estimator, martingale central limit theorem, semimartingales, counting processes, diffusion processes, censored survival data.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ABSTRACT: Consider the simimartingale regression model

$$X(t) = X(0) + \int_0^t Y(s)\alpha(s, Z(s))ds + M(t),$$

Where Y, Z are observable convariate processes, α is a (deterministic) function of both, time and the covariate process Z, and M is a square integrable martingale. Under the assumption

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

that i.i.d. copies of $X < Y < Z$ are observed continuously over a finite time interval, inference for the function $\alpha(t,z)$ is investigated. An estimator $\hat{A}$ for the time integrated $\alpha(t,z)$ and a kernel estimator of $\alpha(t,z)$ itself are introduced. For X a counting process, $\hat{A}$ reduces to the Nelson- Aalen estimator when Z is not present in the model. Various form of consistency are proved, rates of convergence and asymptotic distributions of the estimators are derived. Asymptotic confidence bands for the time integrated $\alpha(t,z)$ and a Kolmogorv-Smirnov-type test of equality of α at different levels of the covariate are given.

END
DATE
FILMED
5-88
DTIC